developersWorks Interviews: Tim Berners-Lee

Originator of the Web and director of the World Wide Web Consortium talks about how far we've come, and about the challenges and opportunities ahead

Level: Introductory

Tim Berners-Lee (timbl@w3.org), Director, World Wide Web Consortium

22 Aug 2006

Podcast with transcript

Tim Berners-Lee is the originator of the World Wide Web and was listed byTime magazine as one of the 100 greatest minds of this century. His inspiring work on the Web set the stage for a world of changes in the way people do business, entertain themselves, exchange ideas, and socialize. In this podcast, Berners-Lee talks about his early history with the Web, opportunities and challenges at present, emerging technologies, and his current project: the semantic Web.

Transcript

developerWorks Interviews: Tim Berners-Lee
Originator of the Web and director of the World Wide Web Consortium
talks about where we've come, and about the challenges and
opportunities ahead
recorded 7-28-2006

LANINGHAM: Welcome to the developerWorks Interviews podcast series.
I'm Scott Laningham, developerWorks podcast editor, and we're IBM's
technical resource for developers, offering a wide range of tools,
code, and education on IBM products and open standards technology.

Today, our guest is Tim Berners-Lee. I'm reading from the
cover flap of the Harper edition of his book, "Weaving the Web, here.
It says, "Tim Berners-Lee, the inventor of the World Wide Web, has been
hailed by Time magazine as one of the 100 greatest minds of this
century. His creation has already changed the way people do business,
entertain themselves, exchange ideas, and socialize with one another.
With new online businesses and communities forming every day, the full
impact of Berners-Lee's grand scheme has yet to be fully known."
Tim Berners Lee is currently the director of the World Wide
Web Consortium, the coordinating body for Web development, and he
occupies the 3COM founder's chair at the M.I.T. laboratory for computer
science and artificial intelligence. He's a recipient of numerous
awards, more than can be mentioned here, including being named a member
of the Royal Society in 2001. For some perspective, Sir Isaac Newton
was the first member of the Royal Society. Tim is currently based in
Cambridge, Massachusetts. Tim, thanks so much for coming on the
podcast today. It's a real honor to talk with you.

BERNERS-LEE: My pleasure.

LANINGHAM: I wanted to ask you at the top here, with the amazing
perspective you have on the last 25 years of technology, especially
that which has grown up around and on top of the Web, do you have a
sense of how you would place it all historically?

BERNERS-LEE: I think, Scott, I'll answer that one in a thousand
years' time, if that's okay.

LANINGHAM: [LAUGHTER] sure.

BERNERS-LEE: That's very difficult...

I think trying to write the history, at the time I think it's very
difficult. Certainly I know during the early days of he Web, for
example, we were watching its exponential growth and during a phase
when the load on our server was actually going up by more or less
exactly a factor of 10 every year.

And we didn't know whether we were going to be looking at history or
not, because when you look at a curve like that, an exponential curve
going up can often just tilt over and crash back down again. And there
were a lot of other projects we've all had which have done that.

So to try to say what the impact on the next few centuries is going to be at this stage I think is really foolish. It's fun to try, but I think there's also so many new things coming up, a lot of them based on the Web, which I think are going to change things again completely.

LANINGHAM: Right. I hear what you're saying. I was just thinking that, without your creative and catalyzing work that spawned the growth of the World Wide Web, we might not be talking about the flattening of the world right now, you know that phrase everybody is using.

BERNERS-LEE: Yes, certainly the fact that we're all connected, the fact that we've got this information space, does change the parameters. It changes the way people live and work. It changes things for good and for bad.

But I think in general it's clear that most bad things come from misunderstanding, and communication is generally the way to resolve misunderstandings, and the Web's a form of communications, so it generally should be good, but I think also we have to watch whether we preserve the stability of the world like we don't want to watch this phenomena like the stock market becoming unstable when it became computerized, for example.

We need to look at the whole society and think, are we actually thinking about what we're doing as we go forward, and are we preserving the really important values that we have in society? Are we keeping it democratic, and open, and so on?

LANINGHAM: That was another question I was going to ask, is are there some negative things that have been unleashed that really demand an attention level from us that's different from things in the past, and I think you've started to speak to that right there. Are there other issues like that surrounding this technology that we need to really be alert to?

BERNERS-LEE: There have been all kinds of things which have...I suppose which have worried us and some of it still do. I think if you go to a person in the street, in fact, the consortium staff once went out into the streets of Cambridge and just interviewed people to see what they thought about the Internet randomly, and I suppose you had to pick one thing, they said, spam, get rid of spam.
[LAUGHTER]

And that in fact for a lot of people makes the e-mail unusable, and it's a ridiculous amount of humanity's time which is being spent just leafing through it and throwing it out.

So there are certainly some things which need really urgent attention in the Internet space. Another thing that happened I suppose during the consortiums, this was the evolution of our process for creating the standards, it started off fairly ad hoc, much more in the start of the Internet engineering task force, people get together, they know each other, and all with a passion for the same thing, get together and produce some technology and roll it out and then sort of create a new market.

Then we found with one particular standard that came through the consortium, that one of the people in the working group turned around and said, oh, right, you're all going to be paying me royalties, by the way, when we're finished. And there was shock and horror. This was during the dot com boom, so this was when things were moving very fast, people were talking about 2.6 Internet years being a year, or was it an Internet year being 2.6 months, things going very fast.

And the whole process stopped for 18 months. It cost us I think $150,000 to find a lawyer who would investigate it and write an opinion. And the opinion was that actually these people would not be paying anybody royalties, that a patent did not apply in that case.

But that sort of sent a shudder through the community, yes, we can all work together producing something. And then somebody could come along and claim that actually they owned the idea.

So that started the consortium on a path of producing a patent policy so that we could make sure that when people come together to make standards they come together understanding that they're trying to produce something which is going to be royalty-free. So it will be generally usable by everybody.

It took several years to actually get that policy agreed, but now it is. So that's the sort of, that was something we may not have anticipated at the beginning, just because the ethos was so much of...where the openness was just, the default, it was the understood thought of the whole Internet community at the beginning.

LANINGHAM: Do you feel like, I've read some of your blog entries about the Net neutrality issue. And do you feel like that subject right now is the same kind of challenge or even more of a challenge than the one you just described in the earlier days of the Web?

BERNERS-LEE: I think the importance in that neutrality is well understood by a lot of people, and I think for example outside the United States I haven't heard any concern about it. If you go to Europe and ask somebody about whether they're concerned about Net neutrality, they say, what? You know, people didn't know the phrase.

It's only that in the US there were some speeches by executives of large telecommunications companies in which they suggested that maybe they should change the whole way they charge for the Internet and block people like Google, so that Google didn't get connection to their customers, quote, for free, unquote. As though Google doesn't pay for its Internet connectivity.

So I think what happened was that put a scare through. And when you look at...when you look at American industry, you know that companies are beholden to their shareholders to try to make whatever money they can. And so where they can find a business plan which they think will produce larger profits in the short term than they may feel obliged to try to pursue that, even though it might be very completely destructive, it would be destructive of the Internet community, the Internet markets and the amazing phenomenon which is built on top of the Internet at the moment.

So I think in fact Congress will understand the concern. There's a huge education process going on at the moment that a lot of people are taking the trouble to explain to American congress what it means, this important concept that if I've connected to the Internet, and if you've connected to the Internet we can then communicate. Nobody else can then start suddenly charging us extra money to building up because they feel that you're making too much money selling me your audio podcast or something.

I think it's really important that the market for Internet connectivity and the market for content are independent markets. It's thus why the Internet has expanded so quickly, because we've had one group of people working on innovations to the Internet infrastructure, and it's more or less since we started, you know, it's gone up, we're talking about, we started off by talking about thousands of bits per second and now we're talking about millions of bits per second, so that's a factor of a thousand increase in speed, which has happened during the time I suppose that the Web was being developed over the last 15 years.

LANINGHAM: Considering that arc of evolution that you're alluding to there a little bit, if we go back, if I understand this bit of history correctly, and around 1980 when you created Enquire utilizing concept of hypertext, and the modern Internet was still unfolding with the first TCP/IP Wide Area Network, was still a few years away, did you have a sort of epiphany with Enquire and that experience along the lines of, wow, I've stumbled on to something here that's far bigger than what can yet be seen or what I think might become of it, but I can tell that this is a sea change moment?

BERNERS-LEE: I don't believe in the sort of eureka moment idea. I think it's a myth. I'm very suspicious that actually Archimedes had been thinking about that problem for a long time.

[LAUGHTER]

And it wasn't that suddenly it came to him. Maybe at some point he felt the pieces had fallen into place. You know, there must have been a point, yes, okay, when he thought, okay, we can lower the crown of the water, and maybe it happened when he was in his bath. But in general, he'd probably, that was backed up probably by an awful lot of thinking about density and volumes and physical structures, which he'd been doing for a long time.

And so Enquire, for example, Enquire was a program which allowed you to make random associations between different things. They were more like note cards. And it was all using much older technology of course, it was using sort of 24 by 80 green terminals.

And it was a step that I'd had a long term interest in the idea that the brain can store random associations between things. You can store a random association between a place, between what you're doing at the moment, for example, if somebody walks in with a particularly flavored cup of coffee, then later on when you remember that flavor you remember what you're doing just now, it will make association between flavors, and running your equipment and conducting a interview, totally otherwise unconnected things.

So computer systems couldn't do that. And so that was a general
interest, which is, and that's related to the ability to make, link
anything to anything, which is a fundamental part of I suppose of the
Web architecture, and then now of the Semantic Web architecture.

But it was no one time when it was just struck with, aha! This will
revolutionize the world. You know? I don't think that happened.

LANINGHAM: I've read some places where Enquire has been described in
the words sort of like a wiki. Is that accurate, and what do you think
about wikis, and is that just another way of doing something
conceptually that was already being done?

BERNERS-LEE: Well, there had always been systems where a lot
of...systems I'd say you could roughly divide into two categories:
those which were, where people created information systems and
created...[in one level] people created it and the other people browsed
it, like sort of CD ROM hypertext based systems. And systems where you
are interacting with the data, things like a calendar. Where when you
look at your calendar, at any point you're browsing through the
calendar, you can put any event in there. Your address book, you can
create a new address, or if you see that somebody's telephone number is
wrong, you can go fix it.

And Enquire was [several nodes], and any point you could be reading
something about something which was documented in the system, you could
press a single key to say, and say link, i want to link this to
something else, or I want to add a new link, add a new data. And
wherever you added new data to it you had to make a link to some
previous data. You could only add new data by linking it on to
something old. So you were actually creating this linked data
structure.

And the original World Wide Web browser of course was also an editor.
I never imagined that anybody would want to write in anchor brackets.
We'd had WYSIWYG editors for a long time. So my function was that
everybody would be able to edit in this space, or different people
would have access rights to different spaces. But I really wanted it
to be a collaborative authoring tool.

And for some reason it didn't really take off that way. And we could
discuss for ages why it didn't. You know, there were browser editors,
maybe the HTML got too complicated for a browser just to be easy.

But I've always felt frustrated that most people don't...didn't have
write access. And wikis and blogs are two areas where suddenly two
sort of genres of online information suddenly allow people to edit, and
they're very widely picked up, and people are very excited about them.

And I think that really for me reinforces the idea that people need to
be creative. They want to be able to record what
they think. They want to be able to, if they see something wrong go
and fix it.

LANINGHAM: Our core audience on developerWorks is obviously software
developers. And I wonder if you have thoughts for that audience

regarding the challenges and opportunities around the Web going forward.

BERNERS-LEE: Well, my passion now is the Semantic Web, of course. And we just, I've got a bunch of undergraduate students at M.I.T. who have been hacking away on a little tabulator program I put together, which is an Ajax program. It's Javascript, it's got an audience library in it, that you can use it to pick up any Semantic Web data, and allow people to browse around through that data, and then analyze it, and map it, and put it on to timelines, and calendars, and maps, and things like that.

And this form of programming where you're programming at the RDF level, I think is going to be the future, because at the moment when you program systems, databases, when you write that sequel you have to be aware of the structure of the database. And the structure of that database is really got a lot of arbitrary design constraints in it, arbitrary design decisions. When somebody designed that database they didn't have tables like this, so they moved from a model of abstract things they wanted to represent into some physical database structure. And when you program this in SQL you have to be aware of that.

And similarly, at an object oriented system you have to be aware of what slots are in each object, and in an XML system you have to be aware of whether we've got an XML file which has got a set of books and in each book there's a shelf, or it's a set of elements for shelves and each element, the shelf has got an element for each book. When you write your X path queries you have to be aware of which way that was done, but that decision is kind of arbitrary.

The nice thing about programming at the RDF level is that you can just say, I'll ask for all the books. You can ask for all the shelves. You can ask for a given shelf whether a book was on it. And you're not worrying so much about the underlying syntax. So I think that we'll be able to write code faster. We'll be able to write application code which
doesn't depend on exactly what sort of XML schema you're using, so that when you find you need to change the schemas you
don't have to rewrite everything.

And the fact that the SPARQL query language has come out now means that all kinds of data can be surfaced as RDF. So I think the exciting thing now is to take existing data sets, whatever they are, look at them and keep the existing systems running but write little SPARQL servers for them just as in the early days with the Web, we kept the existing documentation systems, ran little HTTP servers for them so that the data was served. You can write it. Just as then we wrote little Perl scripts and things to convert the data into HTML, now we can just do the same thing, write these little PHP scripts, or Perl scripts, or whatever is your favorite language, Python say, to convert the data when the request comes in for [UI] convert the data, send out some data in RDF, giving data about that particular URI even if it's a URI for an abstract, some abstract concept.

LANINGHAM: Do you get excited at all about emerging Internet technologies as described, you know, Ajax and things like that? Or are you thinking most of the time at such a conceptual level and a solution

level that you don't really feel like you have time to get too worked
up over those things?

BERNERS-LEE: Oh, this tabulator project, the RDF tabulator, I
coded it up in Ajax. So yes, I'm very...I find that Ajax is the....
It's the one computing platform where everybody who has a Web browser
has got that platform. So the nice thing about it is when you do code
up an Ajax implementation then other people can take it and play with
it. And I think that's one of the things that's so exciting.

Also it's actually got a really powerful, you've got a really powerful
user interface toolkit with the HTML DOM and also with the SVG DOM, so
I think scalable back to graphics, Ajax applications are going to be
really fun in the future.

And then when you've got an RDF and SPARQL library which gives you
access to the Web of data, then that gives you extremely deep
programmatic access both to the user and to the data underneath so that
application becomes something relevant. So yes, I get excited about
things which allow people to have fun and make progress and prototype
new ideas.

LANINGHAM: You know, with Web 2.0, a common explanation out there is
Web 1.0 was about connecting computers and making information
available; and Web 2 is about connecting people and facilitating new
kinds of collaboration. Is that how you see Web 2.0?

BERNERS-LEE: Totally not. Web 1.0 was all about connecting people.
It was an interactive space, and I think Web 2.0 is of course a piece
of jargon, nobody even knows what it means. If Web 2.0 for you is
blogs and wikis, then that is people to people. But that was what the
Web was supposed to be all along.

And in fact, you know, this Web 2.0, quote, it means using the
standards which have been produced by all these people working on Web
1.0. It means using the document object model, it means for HTML and
SVG and so on, it's using HTTP, so it's building stuff using the Web
standards, plus Java script of course.

So Web 2.0 for some people it means moving some of the thinking client
side so making it more immediate, but the idea of the Web as
interaction between people is really what the Web is. That was what it
was designed to be as a collaborative space where people can interact.

Now, I really like the idea of people building things in hypertext, the
sort of a common hypertext space to explain what the common
understanding is and thus capturing all the ideas which led to a given
position. I think that's really important. And I think that blogs and
wikis are two things which are fun, I think they've taken off partly
because they do a lot of the management of the navigation for you and
allow you to add content yourself.

But I think there will be a whole lot more things like that to come,
different sorts of ways in which people will be able to work together.

The semantic wikis are very interesting. These are wikis in which
people can add data and then that data can then be

surfaced and sliced and diced using all kinds of different semantic Web
tools, so that's why it's exciting the way people, things are going,
but I think there are lots of new things in that vein that we have yet
to invent.

LANINGHAM: I have one that I really struggle with, and I talk to co
workers about, one of the tools of the Internet, and that's e-mail, and
how the e-mail client is kind of like the communications desktop or
"the" desktop in the office for so many people.

And I feel like it's archaic in a way that it's a long list of stuff
sorted by date and time or sender, and we waste so much time sorting,
and it's almost like an addiction, you can't pull away from it. Do you
see some kind of paradigm shift coming that will really change the way
we communicate, and that e-mail won't be so dominant in our lives?

BERNERS-LEE: E-mail is interesting, we can't live with it and you
can't live without it [LAUGHTER]. As I said, first of all, spam is
clearly something which needs to be fixed. But if you try to fix e-
mail you'll often end up trying to fix life. In the Web consortium we
use the Web, we live in the Web. So a lot of the things other people
do, back to changing e-mail, we do by editing Web documents. And
sometimes that can be a lot easier. You just have shared editing,
and so documents don't move around, and because they're linked together
we don't change the URIs, that means you can find things. You don't
have to keep running back for e-mails.

So I think one of the ways in which within the Consortium we work
differently is that we have an editable Web. We also use e-mail, and
of course they connect, because every e-mail which is sent to a public
list, or a central list in general, is archived and becomes a Web page.

What's funny is that e-mail isn't completely integrated into this
hypertext link space. E-mail messages have message IDs on them, but if
you click on a message ID there is in fact a message ID, MID:URI space.
If you click on one of those not a lot of systems know how to, what to
do with it. Actually the W3C archive system does have a way of looking
at message IDs, you can give it a message ID and it will give you back
an e-mail if it's got it. But most other systems, a lot of other
systems don't. People aren't used to using them, so they've got these
URIs but they're not actually used, so you can't, often you can't
follow through the threads very easily because a lot of mailers don't
give the links back which allow you to create the thread.

There are some ways in which I think e-mail could be improved. But on
the other hand, the fact that a lot of people want you to do things is
a fact of life.

LANINGHAM: Right [LAUGHTER].

BERNERS-LEE: And organizing the things that you're.... I get a
lot of e mail, and people get...and some of it is from people that I
promised to do things for, and some of it is from random people I don't
know asking for things.

And sorting that and deciding on your priorities, how you're going to spend the next 10 minutes is always a challenge, and that's just part of the art of living your life, I suppose.

LANINGHAM: Can I ask you a wrap-up question by just saying, what are your long term hopes for the Internet and for the Web and this amazing invention that we're all using now? Where do you see its place in our lives in the coming decades?

BERNERS-LEE: Well, the Internet was a really nice clean platform on which I could develop the Web, and I hope the Web will continue to be a nice clean platform on which people keep developing other things.

I hope the Semantic Web will take off so that the data basically all the data which is out there which you have access to, to the Web pages, will now be available as data so you can treat it as data. There will be lots of very exciting applications built on that. And we're starting to see that now, but it really is, you know, we're seriously into the exponential growth of the Semantic Web right now, and that's very exciting.

I think it too will be again a platform on which lots of other things to do will happen. It will be basically used for data interoperability and for data integration across companies and between companies and between applications. But it will then provide a basis for all kinds of exciting new systems built on which will be able to make use of that data.

In general, I hope that we as humanity can learn to use this information space to understand each other, that we can form ourselves into groups in lots of interesting ways so that between us that sort of just tangled web of human groups spans the world and makes that so that it's not...you aren't too many clicks across the social Web from any one person to any other one person, so that start really pulling together in the world and solving the huge challenges which we've got without being distracted by fighting each other.

So I suppose scientific progress and world peace, and the things that we've all got in the back of our minds, have always got to drive us. And I think anybody in fact who does Web development, when you ask them about what they're doing for a few minutes they'll end up back at those fundamental human desires.

LANINGHAM: A wonderful vision, a worthwhile vision. Tim, this has been a real pleasure and honor. Thanks so much for taking time to talk with us today.

BERNERS-LEE: Scott, you're very welcome.

LANINGHAM: Our guest has been Tim Berners-Lee, inventor of the World Wide Web, director of the World Wide Web Consortium, and the 3Com Founders chair at the MIT Laboratory for Computer Science and Artificial Intelligence. Find out more about developerWorks podcasts at ibm.com/developerworks/podcast.

For everyone at developerWorks, I'm Scott Laningham.

Talk to you next time.

[END OF SEGMENT]